

The opinion in support of the decision being entered today was not written for publication and is not binding precedent of the Board.

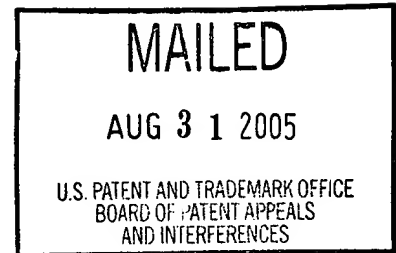
UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

Ex parte FREDERICK J. DAMERAU and DAVID E. JOHNSON

Appeal No. 2005-2260
Application No. 09/605,709

ON BRIEF



Before KRASS, GROSS and BLANKENSHIP, Administrative Patent Judges.

KRASS, Administrative Patent Judge.

Decision On Appeal

This is a decision on appeal from twice-rejected claims 1-6.

The invention pertains to natural language systems and, more particularly, to an automated method for setting up a Web-based conversational natural language interface.

Appeal No. 2005-2260
Application No. 09/605,709

Representative independent claim 1 is reproduced as follows:

1. An automated method for setting up a natural language interface in a Web site comprising the steps of:

defining a hierarchy of topics into which individual documents or Web pages can be classified;

generating a keyword index for those documents; and

for each topic in the hierarchy, associating a set of n-grams to a topic in the topic hierarchy, which set of n-grams is distinctive to that topic and wherein the n-grams maybe sparse or non-sparse n-grams.

The examiner relies on the following reference:

Sarukkai et al. (Sarukkai) 5,819,220 Oct. 6, 1998

Claims 1-6 stand rejected under 35 U.S.C. § 102(b) as anticipated by Sarukkai.

Reference is made to the briefs and answer for the respective positions of appellants and the examiner.

OPINION

A rejection for anticipation under section 102 requires that the four corners of a single prior art document describe every element of the claimed invention, either expressly or inherently, such that a person of ordinary skill in the art could practice the invention without undue experimentation. In re Paulsen, 30 F.3d 1475, 1478-79, 31 USPQ2d 1671, 1673 (Fed. Cir. 1994).

Appeal No. 2005-2260
Application No. 09/605,709

With regard to independent claim 1, the examiner identifies each and every claimed step in Sarukkai as follows:

Column 3, line 56, through column 4, line 7, "in the context of speech interfaces to the web, the invention dynamically makes use of information provided by links in a document or in the current page of a source document being viewed" is said to disclose the claimed "automated method for setting up a natural language interface in a Web site."

Column 7, lines 17-60, and Table 1 therein, is said to disclose the claimed "defining a hierarchy of topics into which individual documents of Web pages can be classified." In particular, Table 1 shows a variety of hierarchical Web links; e.g., <http://www.cs.rochester.edu/> and <http://www.cs.rochester.edu/pub.>

Column 7, lines 17-60, viz., "the information shown in the table was extracted automatically by a simple parsing JAVA program shown in Appendix 1. The set of words constituting the link referent can constitute a web triggered word set, and it would make sense to bias the speech recognition search towards this set of words since it is likely that the user will utter them" is also said to disclose the claimed "generating a keyword index for those documents."

Finally, the examiner points to column 9, lines 17-24, and column 10, lines 16-24, viz., "the concept of extracting web-triggered word set information depending on the context of the web pages recently viewed can also be implemented in other methods. One method would be to appropriately smooth/re-estimate n-gram language model scores using the HTML sources of the documents recently viewed" for a disclosure of the claimed "for each topic in the hierarchy, associating a set of n-grams to a topic in the topic hierarchy, which set of n-grams is distinctive to that topic and wherein the n-grams maybe sparse or non-sparse n-grams."

By pointing out where each and every claimed step can be found in the applied reference, and making a reasonable explanation as to how each of the claimed steps is considered to be disclosed in the reference, in our view, the examiner has set forth a prima facie case of anticipation.

The burden then shifted to appellants to show, if they can, error in the examiner's rationale.

Appellants offer a slew of arguments. Beginning at page 8 of the supplemental brief, appellants offer an explanation of how Sarukkai deals with a voice activated browser while the instant invention "requires a taxonomy of topics for a collection of

Appeal No. 2005-2260
Application No. 09/605,709

documents, assumed to be associated with URLs, and a set of classification rules for each topic," etc.

Such arguments are not persuasive since they do not relate to any specific claim language. While the instant invention may differ from that disclosed by the reference, appellants must point to some specific claim language which is alleged to distinguish over the reference.

At page 9 of the supplemental brief, appellants argue that the term "sparse n-gram," as that term is defined in the instant claims and specification, indicates sequences of tokens or words from the text where the tokens or words may or may not have other words between them. Appellants attempt to distinguish this term over Sarukkai's n-gram which, allege appellants, "means a sequence of tokens that are assigned probabilities within the context of a speech recognition system language model, which is irrelevant to the claimed invention."

We have reviewed page 3, lines 15-27, of the original instant specification, where, allege appellants, there is a definition of sparse and non-sparse n-grams. While that portion of the original specification does explain how gaps are permitted between words making up an n-gram, we find nothing therein offering a "definition" of sparse and non-sparse n-grams.

Appeal No. 2005-2260
Application No. 09/605,709

Moreover, since claim 1 calls for a "set of n-grams" and Sarukkai does disclose such, as pointed out by the examiner, albeit possibly differing from that of the *disclosed* invention, we are not persuaded by appellants' argument that the claimed "n-grams" are somehow different from those disclosed by the reference.

We cannot read limitations from the specification into the application claims. In re Winkhaus, 527 F.2d 637, 188 USPQ 129 (CCPA 1975).

Moreover, as indicated by the examiner, since claim 1 recites, "wherein the n-grams maybe sparse or non-sparse n-grams," the language covers both types of n-grams. Accordingly, even if appellants are correct that Sarukkai does not teach "sparse n-grams," then the reference must teach non-sparse n-grams, still meeting the instant claim language. Moreover, the examiner identifies column 7, lines 65-66, of Sarukkai, a "set of words selectively extracted from the web page source that is being currently displayed by the browser," as the teaching of a "sparse n-gram." Appellants have offered nothing that convinces us that this language of Sarukkai cannot be read as the claimed "sparse n-gram."

Appellants argue, at page 10 of the supplemental brief, that "Sarukkai does not mention using a taxonomy of topics let alone

inducing a taxonomy." However, such argument is not persuasive because the claims do not specifically require any such "taxonomy of topics." Moreover, the examiner indicates, reasonably, in our view, at page 8 of the answer, that Sarukkai's example, in Table 1, of a CS department home page at University of Rochester, and other related topics in a hierarchical manner, is a teaching of using a "taxonomy of topics."

As for appellants' argument that Sarukkai uses n-grams solely for speech recognition, we, again, find ourselves in agreement with the examiner that the reference does teach the use of an n-gram language model (which appellants themselves agree with-see page 11 of the supplemental brief) wherein a web-triggered word set is extracted from a web page source that is being currently displayed from the browser for set up of a natural language interface. Accordingly, we do not find persuasive appellants' argument that the n-grams of the instant invention, created from documents to be searched, are used for very different purposes compared to the n-grams of Sarukkai.

At page 14 of the supplemental brief, appellants argue that Sarukkai does not have a hierarchy of topics, as claimed. Even appellants agree that such a hierarchy is well known (page 14-supplemental brief) but, more importantly, we find it clear from

Sarukkai's Table 1 that the reference certainly does provide for such a hierarchy of topics. Then further on down the page, appellants argue that "whether or not there is a topic hierarchy implicit in Table 1, Sarukkai makes absolutely no explicit use of that information." Presumably, the "use" referred to by appellants is directed to the claimed, "associating a set of n-grams to a topic in the topic hierarchy..." However, as explained by the examiner in the rationale for the rejection, Sarukkai does teach such an association at column 9, lines 17-24, and this does appear to be the case.

In the reply brief, appellants attempt to make a distinction between Sarukkai's finding a textual representation that matches a spoken representation and the sparse n-grams claimed by appellants. We are unpersuaded as the distinctions attempted to be made by appellants have no basis in the *claim* language. Appellants would have us read too much into the claimed term "sparse...n-grams" and we find appellants' interpretation to be overly restrictive.

At page 15 of the supplemental brief, appellants argue the "generating a keyword index..." limitation. Specifically, they argue that Sarukkai "simply does not deal with indexing documents, where the index is to be used for a document search,"

but "only deals with extracting words from documents to bias an acoustic or language model of a speech or voice recognition system." As broadly claimed, we fail to see why Sarukkai's extraction of words from documents cannot be broadly interpreted as "generating a keyword index for those documents." Since the set of words "can constitute a web triggered word set" (column 7, line 20, of Sarukkai), this can be fairly interpreted as the generation of a "keyword index" for those documents. There is nothing in claim 1 which further defines or limits the keyword index in any manner.

Accordingly, since we find none of appellants' arguments anent claim 1 to be persuasive of any error in the examiner's rationale, we will sustain the rejection of claim 1 under 35 U.S.C. § 102(b).

With regard to claim 2, appellants argue that the details of the generating step, i.e., "the step of extracting sparse n-grams of keywords for each group of pages in the topic hierarchy" is not disclosed by Sarukkai. Appellants assert that the examiner's reliance on column 9, lines 19-22, and column 10, lines 16-24 ("n-gram language model score using the HTML sources of the documents recently viewed") is "simply absurd" (supplemental brief-page 16) because the generation of an ordinary search

Appeal No. 2005-2260
Application No. 09/605,709

keyword index has nothing to do with "n-gram language model score..."

The examiner's only response is to point to the cited portion of Sarukkai and state that the examiner "believes that Sarukkai's statement is correct. Sarukkai's invention better explained in his specification how n-gram language model score related with the keyword index" (answer-page 12).

We have reviewed the cited portions of Sarukkai and the examiner's response and we conclude that the examiner has not established a prima facie case of anticipation with regard to claim 2. The claim specifically says that the step of generating a keyword index must comprise the step of "extracting sparse n-grams of keywords for each group of pages in the topic hierarchy." The examiner has utterly failed to show how this is taught by the reference. While Sarukkai clearly extracts web-triggered word set information, we find no indication therein that this equates to extracting sparse n-grams of keywords for each group of pages in the topic hierarchy. If the examiner believes that this concept and how n-gram language model score is related to the keyword index are "better explained" in other portions of Sarukkai, the examiner should have specifically cited

Appeal No. 2005-2260
Application No. 09/605,709

the portions of Sarukkai's disclosure which are relied upon for the rejection.

We will not sustain the rejection of claim 2 under 35 U.S.C. § 102(b).

With regard to claim 3, appellants argue that the step of "optionally reviewing and editing the keyword index" is not taught by Sarukkai and that the examiner's reliance on column 6, lines 36-39 ("[m]odify the appropriate language Model and/or acoustic model parameters dynamically in step 34, using the selected word-set list (see step 32), to be used during the speech recognition search process") is misplaced. In particular, appellants assert that the claimed review and possible modification is manual but that is not the case with parameters of language and/or acoustic models in Sarukkai.

Since the examiner relies on a portion of Sarukkai which does disclose the modification of models using the selected word list, and the word list is the keyword index, it appears reasonable to us to find the limitations of instant claim 3 taught by Sarukkai. Appellants' sole argument is based on their modification being "manual" but this is an argument not based on any limitation appearing in the claim. Accordingly, it is not a persuasive argument for patentability.

Thus, we will sustain the rejection of claim 3 under 35 U.S.C. § 102(b).

With regard to claim 4, the arguments are similar to the arguments presented as per independent claim 1, supra. Accordingly, for similar reasons, we will sustain the rejection of claim 4 under 35 U.S.C. § 102(b). The examiner explains that Sarukkai creates rules, via equation (3), in column 8, and this specific assertion is not argued by appellants, other than to generally deny that Sarukkai teaches the creation of rules.

Turning to claim 5, this claim further limits the creation of rules step, in having this creation performed "automatically" and further comprising the optional step of "manually editing the rules." Since the examiner has not identified any portion of the Sarukkai disclosure for automatically creating rules and optionally manually editing the rules, we will not sustain the rejection of claim 5 under 35 U.S.C. § 102(b).

As per claim 6, the examiner relies on the same rationale for rejecting claim 2. While claim 2 and claim 6 appear to be directed to different limitations, with claim 6 reciting the step of "converting the set of n-grams to classification rules," both appellants and the examiner appear content with letting claim 6 stand or fall with claim 2. Since we have not sustained the

Appeal No. 2005-2260
Application No. 09/605,709

rejection of claim 2, and, further, since the examiner has not identified where, in Sarukkai, is a disclosure of a conversion of a set of n-grams to classification rules, we will not sustain the rejection of claim 6 under 35 U.S.C. § 102(b).

We have sustained the rejection of claims 1, 3, and 4 under 35 U.S.C. § 102(b), but we have not sustained the rejection of claims 2, 5, and 6 under 35 U.S.C. § 102(b).

Accordingly, the examiner's decision is affirmed-in-part.

No time period for taking any subsequent action in connection with this appeal may be extended under 37 CFR § 1.136(a)(1)(iv).

AFFIRMED-IN-PART



ERROL A. KRASS)
Administrative Patent Judge)



ANITA PELLMAN GROSS)
Administrative Patent Judge)

BOARD OF PATENT
APPEALS
AND
INTERFERENCES



HOWARD B. BLANKENSHIP)
Administrative Patent Judge)

EK/rwk

Appeal No. 2005-2260
Application No. 09/605,709

WHITHAM, CURTIS & CHRISTOFFERSON, P.C.
11491 SUNSET HILLS ROAD
SUITE 340
RESTON, VA 20190